

Participant's interest :

1. Reliability of publishers data: bibliographic rules and Onix 3.0 publishers data flows.
2. Open Data aggregation on FRBR work model : Proof of concept "Du livre au film".
3. Features of book information for building a data warehouse

Who is Electre?

Electre references the books written in French by creating records containing bibliographic and commercial data (70.000 books per year).

Electre also publishes a weekly serial "Livres Hebdo", professional publication about the book industry including a large scale of information, articles, sales scores, literature awards, medias agenda...

Since 2010, we catalog books with a suitable FRBR (Functional Requirements for Bibliographic Records) model taking into account the work, expression and manifestation level.

Speakers : Joëlle Aernoudt, François Thévenet, Pierre Boudigues

Who is Antidot?

For over 12 years, Antidot has designed solutions for information access and searching, providing organizations and individuals with all the information they need in order to understand, decide and take action: search engines, capture systems, data structuring and enrichment, information navigation tools, collaboration tools, monitoring tools. Antidot has developed a strategic vision, encompassing the entire field of information searching access, and driven by experience gained in a wide range of projects successfully completed with customers and partners. We developed a data Warehouse with Electre, containing all their data.

Speaker : Gautier Poupeau

Point of view

Metadata is a key feature of book distribution workflows in general, and e-books in particular. Traditional players in the book industry have to take into account the production workflows, quality and scope of their metadata, in order to be able to keep a leading role in the digitisation process unlike what happened for the music industry.

This issue has to be addressed when designing the digital publishing workflow. Metadata management happens at every step of the process and involves every player (publishers, book-sellers, librarians...) each at their own level. Their successful collaboration relies on the use of standards, identifiers and vocabularies, all required in order to reach the necessary interoperability level for exchanging, linking, and using the data they produce.

We would like to share our common expertise in this area, by presenting the result of our work and the various obstacles that have to be tackled when dealing with this issue. Our focus will be on 3 major topics :

- 1) Features of metadata in the book industry
- 2) Sharing and aggregating data in the book industry : do's and don'ts
- 3) Linking book data on the world wide web : example of added value and use cases.

1) Features of metadata in the book industry

Traditional bibliographic records are no longer sufficient to provide useful metadata for the digital world. New models like FRBR, centred on the notion of Work and links between entities, are required.

Several types of metadata go beyond the traditional bibliographic description :

- metadata related to the audience and success of the book like critics (newspaper articles, awards, media events) and user reviews (social networks, comments)
- metadata related to the content (places, stories, characters...)
- metadata related to the author (biography, book signing events...).

2) Sharing and aggregating data in the book industry : dos and don'ts

The different types of metadata mentioned above are not created at the same time and rely on different producers at different steps in the workflow. Aggregating and linking these data requires the common use of standards and identifiers (ISBN, ISTC...).

The example of Electre's datawarehouse using Semantic Web technologies shows the downfalls of current practice in the book industry in France and demonstrates the gap that has to be bridged in order to successfully combine and use the data at a global level.

3) Linking book data on the world wide web : example of added value and use cases

It is important that the data thus aggregated and combined is shared on the web outside the book industry, so that external users can take advantage of highly structured information provided by the traditional players in the field. Then book data producers can link their data to other datasets already available on the web in order to create new services with real added value. We will demonstrate this topic with our prototype “du livre au film” (books and movies).

It is necessary to build a common ecosystem for producers and users in the digital publishing workflow, including a set of standards and identifiers. Semantic web and linked data standards provide a suitable framework for linking successfully different types of data from different producers and silos. The business model for metadata distribution still has to be explored in this perspective.